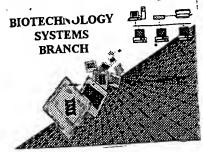# RAW SEQUENCE LISTING
ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: ___09/686,673___

Source: ___OIPE___

Date Processed by STIC: ___10/27/2000___

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.
PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:
1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.
PATENTIN 2.1 e-mail help: patin21help@uspto.gov or phone 703-306-4119 (R. Wax)
PATENTIN 3.0 e-mail help: patin30help@uspto.gov or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE **CHECKER VERSION 3.0 PROGRAM**, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW:

Checker Version 3.0

The Checker Version 3.0 application is a state-of the-art Windows based software program employing a logical and intuitive user-interface to check whether a sequence listing is in compliance with format and content rules. Checker Version 3.0 works for sequence listings generated for the original version of 37 CFR §§1.821 – 1.825 effective October 1, 1990 (old rules) and the revised version (new rules) effective July 1, 1998 as well as World Intellectual Property Organization (WIPO) Standard ST.25.
Checker Version 3.0 replaces the previous DOS-based version of Checker, and is Y2K-compliant. Checker allows public users to check sequence listings in Computer Readable form (CRF) before submitting them to the United States Patent and Trademark Office (USPTO). Use of Checker prior to filing the sequence listing is expected to result in fewer errored sequence listings, thus saving time and money.

Checker Version 3.0 can be down loaded from the USPTO website at the following address:
http://www.uspto.gov/web/offices/pac/checker

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/686,673

DATE: 10/27/2000
TIME: 08:22:54

Input Set : A:\10448-088001Seqlist.ST25.txt
Output Set: N:\CRF3\10272000\I686673.raw

3 <110> APPLICANT: Weich, Nadine
5 <120> TITLE OF INVENTION: 8843, A NOVEL HUMAN DUAL SPECIFICITY PHOSPHATASE FAMILY MEMBER
7 <130> FILE REFERENCE: 10448-088001
9 <140> CURRENT APPLICATION NUMBER: US/09/686,673
9 <141> CURRENT FILING DATE: 2000-10-11
9 <160> NUMBER OF SEQ ID NOS: 8
11 <170> SOFTWARE: PatentIn version 3.0
13 <210> SEQ ID NO: 1
14 <211> LENGTH: 839
15 <212> TYPE: DNA
16 <213> ORGANISM: Homo sapiens
18 <220> FEATURE:
19 <221> NAME/KEY: CDS
20 <222> LOCATION: (44)..(646)
22 <400> SEQUENCE: 1

```
23 cgcgagcgcg ggggccgacg ggtcgccgct gcgccgggcc ggg atg gcg gcc acc          55
                                                    Met Ala Ala Thr
                                                     1
27 gcg ctg ctg gag gcc ggc ctg gcg cgg gtg ctc ttc tac ccg acg ctg         103
28 Ala Leu Leu Glu Ala Gly Leu Ala Arg Val Leu Phe Tyr Pro Thr Leu
                    10                  15                  20
31 ctc tac acc ctg ttc cgc ggg aag gtg ccg ggt cgg gcg cac cgg gac         151
32 Leu Tyr Thr Leu Phe Arg Gly Lys Val Pro Gly Arg Ala His Arg Asp
                25                  30                  35
35 tgg tac cac cgc atc gac ccc acc gtg ctg ctg ggc gcg ctg ccg ttg         199
36 Trp Tyr His Arg Ile Asp Pro Thr Val Leu Leu Gly Ala Leu Pro Leu
            40                  45                  50
39 cgg agc ttg acg cgc cag ctg gta cag gac gag aac gtg cgc ggg gtg         247
40 Arg Ser Leu Thr Arg Gln Leu Val Gln Asp Glu Asn Val Arg Gly Val
        55                  60                  65
43 atc acc atg aac gag gag tac gag acg agg ttc ctg tgc aac tct tca         295
44 Ile Thr Met Asn Glu Glu Tyr Glu Thr Arg Phe Leu Cys Asn Ser Ser
    70                  75                  80
47 cag gag tgg aag aga cta gga gtc gag cag ctg cgg ctc agc aca gta         343
48 Gln Glu Trp Lys Arg Leu Gly Val Glu Gln Leu Arg Leu Ser Thr Val
85                  90                  95                  100
51 gac atg act ggg atc ccc acc ttg gac aac ctc cag aag gga gtc caa         391
52 Asp Met Thr Gly Ile Pro Thr Leu Asp Asn Leu Gln Lys Gly Val Gln
                105                 110                 115
55 ttt gct ctc aag tac cag tcg ctg ggc cag tgt gtt tac gtg cat tgt         439
56 Phe Ala Leu Lys Tyr Gln Ser Leu Gly Gln Cys Val Tyr Val His Cys
            120                 125                 130
59 aag gct ggg cgc tcc agg agt gcc act atg gtg gca gca tac ctg att         487
60 Lys Ala Gly Arg Ser Arg Ser Ala Thr Met Val Ala Ala Tyr Leu Ile
        135                 140                 145
63 cag gtg cac aaa tgg agt cca gag gag gct gta aga gcc atc gcc aag         535
64 Gln Val His Lys Trp Ser Pro Glu Glu Ala Val Arg Ala Ile Ala Lys
```

```
65         150                          155                    160                              583
67 atc cgg tca tac atc cac atc agg cct ggc cag ctg gat gtt ctt aaa
68 Ile Arg Ser Tyr Ile His Ile Arg Pro Gly Gln Leu Asp Val Leu Lys
69 165                       170                    175                    180               631
71 gag ttc cac aag cag att act gca cgg gca aca aag gat ggg act ttt
72 Glu Phe His Lys Gln Ile Thr Ala Arg Ala Thr Lys Asp Gly Thr Phe
73                 185                    190                    195
75 gtc att tca aag aca tgatgtatgg ggattagaaa gaactcaaga cactcctgct              686
76 Val Ile Ser Lys Thr
77                 200
79 tgatacagaa caaaaagagc ttaacaggac caacagggct taagcccaga cttgacgtaa           746
81 cagaaatgtg ccaataggta ataggtaatt tttctttctc tgacttgttt tgttttcttg           806
83 aaataacact gttgtgtggc tagaaaaaaa aaa                                        839
86 <210> SEQ ID NO: 2
87 <211> LENGTH: 201
88 <212> TYPE: PRT
89 <213> ORGANISM: Homo sapiens
91 <400> SEQUENCE: 2
93 Met Ala Ala Thr Ala Leu Leu Glu Ala Gly Leu Ala Arg Val Leu Phe
94 1               5                   10                      15
97 Tyr Pro Thr Leu Leu Tyr Thr Leu Phe Arg Gly Lys Val Pro Gly Arg
98               20                  25                      30
101 Ala His Arg Asp Trp Tyr His Arg Ile Asp Pro Thr Val Leu Leu Gly
102              35                  40                      45
105 Ala Leu Pro Leu Arg Ser Leu Thr Arg Gln Leu Val Gln Asp Glu Asn
106        50                  55                      60
109 Val Arg Gly Val Ile Thr Met Asn Glu Glu Tyr Glu Thr Arg Phe Leu
110 65                 70                  75                      80
113 Cys Asn Ser Ser Gln Glu Trp Lys Arg Leu Gly Val Glu Gln Leu Arg
114              85                  90                      95
117 Leu Ser Thr Val Asp Met Thr Gly Ile Pro Thr Leu Asp Asn Leu Gln
118              100                 105                     110
121 Lys Gly Val Gln Phe Ala Leu Lys Tyr Gln Ser Leu Gly Gln Cys Val
122        115                 120                     125
125 Tyr Val His Cys Lys Ala Gly Arg Ser Arg Ser Ala Thr Met Val Ala
126        130                 135                     140
129 Ala Tyr Leu Ile Gln Val His Lys Trp Ser Pro Glu Glu Ala Val Arg
130 145                 150                 155                     160
133 Ala Ile Ala Lys Ile Arg Ser Tyr Ile His Ile Arg Pro Gly Gln Leu
134              165                 170                     175
137 Asp Val Leu Lys Glu Phe His Lys Gln Ile Thr Ala Arg Ala Thr Lys
138              180                 185                     190
141 Asp Gly Thr Phe Val Ile Ser Lys Thr
142              195                 200
145 <210> SEQ ID NO: 3
146 <211> LENGTH: 606
147 <212> TYPE: DNA
148 <213> ORGANISM: Homo sapiens
150 <400> SEQUENCE: 3
```

```
151 atggcggcca ccgcgctgct ggaggccggc ctggcgcggg tgctcttcta cccgacgctg    60
153 ctctacaccc tgttccgcgg gaaggtgccg ggtcgggcgc accgggactg gtaccaccgc   120
155 atcgacccca ccgtgctgct gggcgcgctg ccgttgcgga gcttgacgcg ccagctggta   180
157 caggacgaga acgtgcgcgg ggtgatcacc atgaacgagg agtacgagac gaggttcctg   240
159 tgcaactctt cacagggagtg gaagagacta ggagtcgagc agctgcggct cagcacagta   300
161 gacatgactg ggatccccac cttggacaac ctccagaagg gagtccaatt tgctctcaag   360
163 taccagtcgc tgggccagtg tgtttacgtg cattgtaagg ctgggcgctc caggagtgcc   420
165 actatggtgg cagcatacct gattcaggtg cacaaatgga gtccagagga ggctgtaaga   480
167 gccatcgcca agatccggtc atacatccac atcaggcctg gccagctgga tgttcttaaa   540
169 gagttccaca agcagattac tgcacgggca acaaaggatg ggactttgt catttcaaag   600
171 acatga                                                             606
```

```
174 <210> SEQ ID NO: 4
175 <211> LENGTH: 173
176 <212> TYPE: PRT
177 <213> ORGANISM: (Artificial/Unknown)
179 <220> FEATURE:
180 <221> NAME/KEY: VARIANT
181 <222> LOCATION: (1)..(173)
182 <223> OTHER INFORMATION: consensus sequence
```

per new Sequence Rules, the only valid responses are:
Unknown OR
Artificial Sequence
or scientific name
(Genus/species)

(global error)

```
185 <400> SEQUENCE: 4
187 Gly Pro Ser Glu Ile Leu Pro His Leu Tyr Leu Gly Ser Tyr Ser Thr
188 1               5                   10                  15
190 Ala Ser Glu Ala Asn Leu Ala Leu Leu Lys Lys Leu Gly Ile Thr His
191             20                  25                  30
193 Val Ile Asn Val Thr Glu Glu Val Pro Asn Pro Phe Glu Leu Asp Lys
194         35                  40                  45
196 Lys Asn Asp Arg His Tyr Thr Asn Ala Tyr Ile Ser Lys Asn Ser Gly
197     50                  55                  60
199 Phe Thr Tyr Leu Gln Ile Pro Asn Val Asp Asp His Ile Tyr Tyr His
200 65                  70                  75                  80
202 Ile Ala Trp Asn His Glu Thr Lys Ile Ser Lys Tyr Phe Asp Glu Ala
203             85                  90                  95
205 Val Asp Phe Ile Asp Asp Ala Arg Gln Lys Gly Gly Lys Val Leu Val
206             100                 105                 110
208 His Cys Gln Ala Gly Ile Ser Arg Ser Ala Thr Leu Ile Ile Ala Tyr
209             115                 120                 125
211 Leu Met Lys Thr Arg Asn Leu Ser Leu Asn Glu Ala Tyr Asp Phe Val
212             130                 135                 140
214 Tyr Val Tyr His Ile Lys Glu Arg Arg Cys Pro Ile Ile Ser Pro Asn
215 145                 150                 155                 160
217 Phe Gly Phe Leu Arg Gln Leu Ile Glu Tyr Glu Arg Lys
218             165                 170
```

```
220 <210> SEQ ID NO: 5
221 <211> LENGTH: 172
222 <212> TYPE: PRT
223 <213> ORGANISM: (Artificial/Unknown)
225 <220> FEATURE:
226 <221> NAME/KEY: VARIANT
227 <222> LOCATION: (1)..(172)
```

also, explain in <223>, if <213> response is
Artificial Sequence or
Unknown

(see next page)

10/27/00

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/686,673

DATE: 10/27/2000
TIME: 08:22:54

Input Set : A:\10448-088001Seqlist.ST25.txt
Output Set: N:\CRF3\10272000\I686673.raw

No Xaa's in this sequence

```
228 <223> OTHER INFORMATION: Xaa = Any Amino Acid
231 <400> SEQUENCE: 5
233 Gly Pro Ser Glu Ile Leu Pro His Leu Tyr Leu Gly Ser Tyr Ser Asp
234 1               5                   10                  15
236 Ala Ser Glu Ala Asn Leu Ala Leu Leu Lys Lys Leu Gly Ile Thr His
237         20                  25                  30
239 Val Ile Asn Val Thr Glu Glu Val Pro Asn Asn Phe Glu Leu Lys Lys
240         35                  40                  45
242 Lys Asn Asp Arg Tyr Tyr Thr Asn Glu Tyr Ile Ser Lys Gly Ser Gly
243         50                  55                  60
245 Phe Thr Tyr Leu Gln Ile Pro Asn Val Asp Asp Ile Tyr Tyr His Ile
246 65                  70                  75                  80
248 Ala Trp Asn Thr Glu Thr Lys Ile Ser Lys Tyr Leu Glu Glu Ala Val
249             85                  90                  95
251 Glu Phe Ile Glu Asp Ala Glu Lys Lys Gly Gly Lys Val Leu Val His
252             100                 105                 110
254 Cys Gln Ala Gly Val Ser Arg Ser Ala Thr Leu Val Ile Ala Tyr Leu
255         115                 120                 125
257 Met Lys Thr Arg Asn Leu Ser Leu Arg Asp Ala Tyr Asp Phe Val Tyr
258         130                 135                 140
260 Val Tyr His Ile Lys Glu Arg Arg Cys Pro Ile Ile Ser Pro Asn Phe
261 145                 150                 155                 160
263 Gly Phe Leu Arg Gln Leu Ile Glu Tyr Glu Arg Lys
264             165                 170

266 <210> SEQ ID NO: 6
267 <211> LENGTH: 13
268 <212> TYPE: PRT
269 <213> ORGANISM: Artificial/Unknown
271 <220> FEATURE:
272 <221> NAME/KEY: VARIANT
273 <222> LOCATION: (1)..(1)
274 <223> OTHER INFORMATION: Xaa = Leu, Ile, Val, Met, or Phe
277 <220> FEATURE:
278 <221> NAME/KEY: VARIANT
279 <222> LOCATION: (1)..(13)
280 <223> OTHER INFORMATION: active site signature
283 <220> FEATURE:
284 <221> NAME/KEY: VARIANT
285 <222> LOCATION: (4)..(9)
286 <223> OTHER INFORMATION: Xaa = Any Amino Acid
289 <220> FEATURE:
290 <221> NAME/KEY: VARIANT
291 <222> LOCATION: (10)..(10)
292 <223> OTHER INFORMATION: Xaa = Ser, Thr, or Cys
295 <220> FEATURE:
296 <221> NAME/KEY: VARIANT
297 <222> LOCATION: (11)..(11)
298 <223> OTHER INFORMATION: Xaa = Ser, Thr, Ala, Gly, or Pro
301 <220> FEATURE:
```

```
302 <221> NAME/KEY: VARIANT
303 <222> LOCATION: (12)..(12)
304 <223> OTHER INFORMATION: Xaa = Any Amino Acid
307 <220> FEATURE:
308 <221> NAME/KEY: VARIANT
309 <222> LOCATION: (13)..(13)
310 <223> OTHER INFORMATION: Xaa = Leu, Ile, Val, Met, Phe, or Tyr
313 <400> SEQUENCE: 6
315 Xaa His Cys Xaa Xaa Gly Xaa Xaa Xaa Xaa Xaa Xaa Xaa
316 1                     5                    10
318 <210> SEQ ID NO: 7
319 <211> LENGTH: 7
320 <212> TYPE: PRT
321 <213> ORGANISM: Artificial/Unknown
323 <220> FEATURE:
324 <221> NAME/KEY: VARIANT
325 <222> LOCATION: (1)..(7)
326 <223> OTHER INFORMATION: Xaa = Any Amino Acid
329 <220> FEATURE:
330 <221> NAME/KEY: VARIANT
331 <222> LOCATION: (1)..(7)
332 <223> OTHER INFORMATION: motif sequence
335 <400> SEQUENCE: 7
337 Cys Xaa Xaa Xaa Xaa Xaa Arg
338 1               5
340 <210> SEQ ID NO: 8
341 <211> LENGTH: 21
342 <212> TYPE: PRT
343 <213> ORGANISM: Artificial/Unknown
345 <220> FEATURE:
346 <221> NAME/KEY: VARIANT
347 <222> LOCATION: (1)..(17)
348 <223> OTHER INFORMATION: Xaa = Any Amino Acid
351 <220> FEATURE:
352 <221> NAME/KEY: VARIANT
353 <222> LOCATION: (20)..(20)
354 <223> OTHER INFORMATION: Xaa = Leu or Ile
357 <400> SEQUENCE: 8
359 Val Xaa Val His Cys Xaa Xaa Gly Xaa Ser Arg Ser Xaa Thr Xaa Xaa
360 1               5                   10                  15
362 Xaa Ala Tyr Xaa Met
363                 20
```

OK

OK ->

Xplain <213> response in <223>

OK

W

VERIFICATION SUMMARY
PATENT APPLICATION: US/09/686,673

Input Set : A:\10448-088001Seqlist.ST25.txt
Output Set: N:\CRF3\10272000\I686673.raw

L:9 M:270 C: Current Application Number differs, Replaced Current Application No
L:9 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:315 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6
L:337 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7
L:359 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8
L:362 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8